Overview

This document describes the OMI SO_2 product (OMSO2) produced from global mode UV measurements of the Ozone Monitoring Instrument (OMI). OMI was launched on July 15, 2004 on the EOS Aura satellite, which is in a sun-synchronous ascending polar orbit with 1:45 pm local equator crossing time. The data collection started on August 17, 2004 (orbit 482) and continues to this day with only minor data gaps. The minimal volcanic SO_2 mass detectable by OMI is about two orders of magnitude smaller than the detection threshold of the legacy Total Ozone Mapping Spectrometer (TOMS) SO_2 data (1978-2005) [Krueger, et al. 1995]. OMI also enables the detection of anthropogenic SO_2 pollution in the lowest part of the atmosphere. This is due to smaller OMI footprint and the use of wavelengths better optimized for separating O_3 from SO_2.

The product file, called a data granule, covers the sunlit portion of the orbit with an approximately 2600 km wide swath. Each swath normally contains approximately 1600 viewing lines along the ground track of the satellite, with each viewing line containing 60 pixels or scenes across the satellite track. Scenes from all viewing lines with the same cross-track scene number are referred to as a row of the OMI swath. During normal operations, 14 or 15 granules are produced daily, providing fully contiguous coverage of the globe. Currently, OMSO2 products are not produced when OMI goes into the "zoom mode" for one day every 452 orbits (~32 days).

Since 25 June 2007 signal suppression (anomaly) has been observed in Level 1B Earth radiance data for scenes in rows 53-54 (0-based). This anomaly is also known as the OMI row anomaly since it affects some particular rows of the CCD detector. It has since expanded to affect more rows. In SO_2 data, the row anomaly manifests itself as positive or negative stripes (discontinuity in SO_2 with cross-track viewing angle). Efforts have been made to flag the affected scenes. SO_2 data fields for scenes determined to have been influenced by the row anomaly have been assigned a large negative fill-value. More information about the OMI row anomaly can be found from KNMI.

For each OMI scene we provide 4 different estimates of the column density of SO_2 in Dobson Units (1DU=2.69 $\cdot 10^{16}$ molecules/cm^2) obtained by making different assumptions about the vertical distribution of the SO_2. However, it is important to note that in most cases the precise vertical distribution of SO_2 is unimportant. The users can use either the SO_2 plume height, or the center of mass altitude (CMA) derived from SO_2 vertical distribution, to interpolate between the 4 values:

- Planetary Boundary Layer (PBL) SO_2 column (**ColumnAmountSO2_PBL**), corresponding to CMA of 0.9 km. **Please check the following section for important updates to the PBL SO_2 data.**

- Lower tropospheric SO_2 column (**ColumnAmountSO2_TRL**), corresponding to CMA of 2.5 km.

- Middle tropospheric SO_2 column, (**ColumnAmountSO2_TRM**), usually produced by volcanic degassing, corresponding to CMA of 7.5 km,

- Upper tropospheric and Stratospheric SO_2 column (**ColumnAmountSO2_STL**), usually produced by explosive volcanic eruption, corresponding to CMA of 17 km.

The accuracy and precision of the derived SO_2 columns vary significantly with the SO_2 CMA and column amount, observational geometry, and slant column ozone. OMI becomes more sensitive to SO_2 above clouds and snow/ice, and less sensitive to SO_2 below clouds. Preliminary error estimates are discussed below (see Data Quality Assessment).

Important Updates to OMI PBL SO_2 Data

The SO_2 data in **ColumnAmountSO2_PBL** are now produced with a completely different retrieval algorithm based on principal component analysis (PCA) of the OMI radiance data [Li et al. 2013]. Previously the OMI PBL SO_2 data were produced using the Band Residual Difference (BRD) algorithm [Krotkov et al 2006]. While the BRD algorithm is sensitive to SO_2 pollution in the PBL, it tends to have large noise and unphysical biases particularly at high latitudes. The PCA algorithm greatly improves the quality of OMI SO_2 retrievals and has been implemented for operational production of the next generation OMI standard SO_2 product. **PBL SO_2 data users who have acquired OMSO2 data prior to October 2014 are strongly encouraged to download and use the new OMSO2 data. All SO_2 data fields ending with "BRD" are obsolete and only used for internal diagnostic purposes.**

Algorithm Description

We use two different algorithms to produce SO_2 column amount data from OMI. The PBL columns are produced using the principal component analysis (PCA) algorithm [Li et al 2013] that are sensitive to pollution near the surface, while TRL, TRM and STL columns are produced with the Linear Fit (LF) algorithm for volcanic SO_2 [Yang et al 2007].

In the PCA algorithm, we apply a principal component analysis technique to radiance data over a presumably SO_2-free region (*e.g.,* the equatorial Pacific). The resulting principal components (PCs) can capture most (> 99.9999%) of measurement-to-measurement variation of the radiances. The PCs are ordered so that the first PC explains the most of variance, the second PC explains the second most of variance, and so on. The first few leading PCs are generally associated with geophysical processes including ozone absorption, surface reflectance, and rotational-Raman scattering effects (RRS, also known as the Ring effect), while the following PCs often have high-frequency features likely originating from measurement noise and detector artifacts such as wavelength shift and stretch. These physical processes and measurement details can cause strong interferences in SO_2 retrievals, and the PCs enable us to appropriately account for them. By fitting a set of n_v PCs (v_i) along with the SO_2 Jacobians, which represents the

sensitivity of the radiances to the SO_2 column ($\partial N/\partial \Omega_{SO_2}$), to the measured Sun-normalized radiances, we can simultaneously obtain estimates of SO_2 column density (Ω_{SO2}) and coefficients of the PCs (***ω***):

$$N\left(\omega, \Omega_{SO_2}\right) = \sum_{i=1}^{n_v} \omega_i v_i + \Omega_{SO_2} \frac{\partial N}{\partial \Omega_{SO_2}}, \qquad (1)$$

Here N is the measured N-value spectrum ($N(\lambda) = -100 \times \log_{10}(I(\lambda)/I_0(\lambda)$, I and I_0 are radiance and irradiance at wavelength λ, respectively) for a given OMI scene. The PCA algorithm shares the same overall physics concept with the widely used Differential Optical Absorption Spectroscopy (DOAS) method, but the data-driven (vs. forward modeling) approach used to account for retrieval interferences reduces modeling uncertainties, enhances computation efficiency, and makes the PCA algorithm much less sensitive to instrument calibration issues. A more detailed discussion of the PCA algorithm can be found in *Li et al*. [2013] and *Joiner et al.* [2013].

For input data, the PCA algorithm uses OMI level 1B (L1B) radiance and irradiance data in the spectral window of 310.5-340 nm, as well as the O_3 column amount (Ω_{O3}) from the OMTO3 product [*Bhartia and Wellemeyer*, 2002]. The spectral window includes the strong SO_2 absorption band at 310.8 nm and minimizes potential interferences due to stray light at shorter wavelengths. To better account for the orbit-to-orbit measurement artifacts and the different characteristics of the 60 rows of the OMI detector, we process data from each row of each orbit separately. Scenes having strong O_3 absorption due to large slant column O_3 (S_{O3} > 1500 DU) are filtered out before PCA, given the much smaller expected SO_2 sensitivity for these scenes. After data filtering, we first conduct PCA on the approximately 900-1300 remaining scenes for an entire row, without screening out polluted areas. Since SO_2 absorption is generally very weak outside of polluted and volcanic-affected areas, it is unlikely for the PC(s) associated with or affected by SO_2 absorption (v_{so2}) to be among the first few leading PCs. A correlation analysis between the PCs and the SO_2 Jacobians is then conducted to determine the number of PCs (n_v) to be included in the fitting. This ensures that n_v is sufficiently small to prevent the inclusion of v_{so2} and collinearity in Eq. 1, and allows reasonable initial estimates of SO_2 (Ω_{SO2_ini}) to be obtained. To maintain computational efficiency, we set an upper limit of 20 for n_v. A second step PCA is then applied to scenes with small Ω_{SO2_ini} (within ±1.5 standard deviations for each orbit/row) to extract a new set of PCs to update Eq. 1, followed by updated retrievals of SO_2. This step is repeated twice, as the changes in the retrieved SO_2 generally become very small within two iterations. The second step PCA and retrievals are carried out separately for three segments of each row: a "tropical" region with S_{O3} < 100 DU + min(S_{O3}), and two regions north and south of it. These regionally derived PCs more closely match the measurements and help reduce retrieval biases.

The SO_2 Jacobians used in the current version of the PCA algorithm are calculated with the VLIDORT radiative transfer code [*Spurr*, 2008]. The calculation assumes the same measurement conditions as those in the BRD algorithm. More specifically, we assume fixed surface albedo (0.05), surface pressure (1013.25 hPa), as well as fixed solar zenith

angle (30°) and viewing zenith angle (0°). For SO_2, a climatological profile over the summertime eastern U.S. are used. For O_3 and temperature, the OMTO3 standard mid-latitude profiles with Ω_{O3} = 325 DU are used. This setup allows direct comparison between the new and old OMI PBL SO_2 data. In the future, we plan to expand the look-up table for SO_2 Jacobians to more realistically account for different measurement conditions.

The LF algorithm uses a recently modified version (Version 8.5) of TOMS total ozone algorithm (OMTO3) [Bhartia and Wellemeyer 2002] as a linearization step to derive an initial estimate of total ozone assuming zero SO_2. (See OMTO3 README file for more detail). The residuals at the 10 wavelengths are then calculated as the difference between the measured and computed N-values using a vector forward model radiative transfer code that accounts for multiple Rayleigh scattering, ozone absorption, Ring effect, and surface reflectivity, but assumes no aerosols. Cloudy scenes are treated as mixture of two opaque Lamberian surfaces, one at the terrain pressure and the other at Radiative Cloud Pressure (RCP) derived using OMI-measured Rotational Raman scattering at around 350 nm (see OMCLDRR README file for more detail). In the presence of SO_2, the residuals contain spectral structures that correlate with the SO_2 absorption cross-section. The residuals also have contributions from errors sources that have not yet been identified. To reduce this interference, a median residual for a sliding group of SO_2-free and cloud-free scenes (OMTO3 radiative cloud fraction < 0.15) covering ±15° latitude along the orbit track is subtracted for each spectral band and cross-track position [Yang et al 2007].

The LF algorithm uses the corrected residuals as input. SO_2 produced by volcanic degassing and eruptions can produce large errors in OMTO3 derived total ozone and can make the retrieval highly non-linear. The linear Fit (LF) algorithm was developed to handle such cases. The LF algorithm minimizes different subsets of residuals by simultaneously adjusting total SO_2, ozone and includes a quadratic polynomial in the spectral fit. The subsets are determined by the process of dropping the shortest wavelength bands one at a time until the 322nm band is reached. The largest SO_2 retrieval is reported as the final estimate. The assumed gaseous vertical profiles correspond to the standard OMTO3 ozone profiles. The SO_2 weighting functions are approximated using OMTO3 layer Efficiency factors in Umkehr layers 0, 1 and 3, for ColumnAmountSO2_TRL, ColumnAmountSO2_TRM, and ColumnAmountSO2_STL data, correspondingly. Treatment of aerosols and clouds is the same as in the OMTO3 algorithm.

Data Quality Assessment

Errors in OMI SO_2 data can arise from both the input radiance/residual data and the SO_2 Jacobians/weighting functions used in retrievals. For the LF algorithm, the "sliding median" empirical residual correction essentially acts as a high-pass filter reducing cross-track and low frequency latitudinal biases, but allowing high frequency (i.e. "scene by

scene") noise in the residuals to propagate into retrieved background SO_2 data. The resulted errors are best described as pseudo-random (i.e. having different systematic and random components depending on spatial and temporal scales) Gaussian-like distribution with a nominal mean of zero. The errors usually reduce much slower than the square root of the number of measurements averaged. The noise in PCA retrievals can also be described in a similar fashion.

We provide separate Quality Flags (QF) for each of the products that are based on SO_2 consistency criteria between the individual wavelength pairs. The OMSO2 scene quality flag is an automatic assessment of the SO_2 values for the corresponding scene by the OMSO2 retrieval algorithm. It is used primarily as an indicator of the validity of the retrieved SO2 values. For detailed information about the OMSO2 quality flag, please consult the OMSO2 file specification).**While the quality flag may provide some information on the usefulness of retrievals, we have found it to be too restrictive and not very useful in its current form.** Preliminary analysis of the QF values has shown that they work best for large volcanic events, but miss many real PBL and low level degassing emissions. Therefore, independent verification of the real SO_2 signal is strongly recommended. **OMSO2 data users are advised to ignore the quality flag in the current version and use other parameters such as solar zenith angle for data filtering, as specified below for the PBL data**. No data filtering is needed for TRM, TRL, and STL data fields. Below are data quality assessments for each SO_2 product after applying the "sliding median" empirical residual correction and ignoring QF (**Note: no sliding median is applied to PBL**). For all products the noise increases with increasing solar zenith angle at high latitudes and in the region of "South Atlantic radiation Anomaly".

ColumnAmountSO2_PBL: As a measurement of retrieval noise, the standard deviation (sigma) for instantaneous field of view (IFOV) is ~0.5 DU over the presumably SO_2-free equatorial Pacific, or about half that of the BRD algorithm. The root mean square (RMS) for IFOV in different latitude bands over the Pacific can be viewed as a measure of both noise and biases in retrievals, and is estimated at ~0.5 DU for regions between 30°S and 30°N, suggesting very small systematic biases in PCA retrievals over the tropics. The IFOV RMS of PCA retrievals increases to ~0.7-0.9 DU for high latitude regions with large slant column O_3, but is still more than a factor of two smaller than that of BRD retrievals. **Data users are advised to use caution when analyzing data from the edges of the OMI swath (rows 0 and 59, 0-based), as they tend to have greater noise. For best data quality, use data from scenes near the center of the swath (rows 4-54, 0-based) with slant column O_3 < 1500 DU. Retrievals for OMI scenes from the descending node of the Aura satellite should not be used.** The PCA retrievals also have a negative bias over some highly reflective surfaces such as certain areas in the Sahara (up to about -0.5 DU in monthly mean). This negative bias is small as compared to the biases in the BRD retrievals, and is expected to be further reduced after the implementation of a more extensive Jacobians lookup table (see below). For cloudy scenes, the BRD algorithm sometimes produces large negative retrievals, a bias that is now eliminated in the PCA retrievals.

The SO_2 retrieval accuracy also depends on the error in the SO_2 Jacobians. This error is systematic and increases with deviation of the observational conditions from those assumed in the Jacobian calculation. SO_2 will likely be overestimated for remote oceanic regions where SO_2 is transported from source regions and likely located at elevated levels above the PBL. Likewise, SO_2 will also be overestimated for scenes with snow/ice and/or clouds. We plan to expand the look-up table for SO_2 Jacobians to more realistically account for different measurement conditions. **Before this improvement, only snow/ice-free scenes with radiative cloud fraction < 0.3 should be used in studies on SO_2 emission sources.** The higher reflectivity scenes can still be used to track long-range transport of sulfur pollution. Finally, there is also a small but noticeable dependence of retrieved SO_2 on the number of PCs included in the fitting. We expect to provide a more complete error estimate in follow-up releases.

ColumnAmountSO2_TRL: Due to increased sensitivity to elevated SO2, the 1 sigma noise in TRL data is estimated at ~0.7 DU under optimal observational conditions in the tropics. The data can be used for cloudy, clear and mixed scenes as well as for elevated terrain. However, the TRL data contain filled values when terrain pressure or RCP is less than ~500hPa. In such cases the cloud blocks most of the SO2. As a result, the SO2 weighting function approaches zero, no LF retrieval is done and the fill value is stored in the output.

ColumnAmountSO2_TRM are optimized for typical volcanic degassing from volcanoes with vents at ~5km altitude or above and emissions from effusive eruptions. The standard deviation of TRM retrievals in background areas is about 0.3 DU at low and mid-latitudes. The cloud-related fill values in TRM data occurs only when the OMI measured cloud top is higher than ~8-10 km. Biases in the TRM retrievals due to latitude and viewing angle are removed to the 0.1 DU level by the median residual background corrections. Both the bias and standard deviations increase with solar zenith angle. **We recommend that the TRM retrievals be used for volcanic degassing cases at all altitudes** because the PBL retrievals are restricted to optimal viewing conditions and TRL data are overestimated for high altitude emissions (>3km). In general, SO_2 releases at altitudes less than ~7.5 km will be underestimated, but these errors can be corrected off-line using the averaging kernel [Yang et al 2007] if the actual SO_2 vertical distribution is known.

Analysis of daily OMSO2 data for degassing volcanoes at high altitude (~5 km) has shown that significant trends in SO_2 burdens, linked to variability of source SO_2 emissions, can be detected [Carn et al., 2008a]. Preliminary surveys of global volcanic OMSO2 data indicate that the current sensitivity of the algorithm permits detection of volcanoes emitting on the order of 103 tons SO_2/day or more in daily data (under optimal viewing conditions). Detection of weaker sources usually requires temporal averaging of the OMSO2 data.

ColumnAmountSO2_STL data are intended for use with explosive volcanic eruptions where the cloud is placed in the upper troposphere or stratosphere (UTLS). At these altitudes the averaging kernel is weakly dependent on altitude, so that differences in

actual cloud height from ~17 km produce only small errors. The biases with latitude and viewing angle are generally less than 0.2 DU. The noise level in background data is about 0.2 DU. This sensitivity has permitted tracking of volcanic SO_2 clouds in the UTLS for great distances from the source [e.g., Carn et al., 2007b, Carn et al., 2008b]. Both the bias and standard deviation increase near the northern terminator, similar to but reduced from the TRM results. Artifacts due to ozone profile errors are reduced from the TRM data by about 30%. One should see no fill values due to cloud screening in the STL data.

The LF algorithm still has large error when it comes to high SO_2 loading cases. The LF algorithm as implemented in the v1.1.6 OMSO2 is expected to provide good retrieval when SO_2 loading is less than ~50 DU. When SO_2 loadings are higher than ~100 DU the LF algorithm underestimates the true SO_2 amount, the higher the loading the larger the underestimation [Yang et al 2007]. Comparisons between total SO_2 burdens calculated using OMSO2 and EP-TOMS SO_2 data for volcanic clouds in the UTLS have shown agreement to within 20% for SO_2 column amounts of <100 DU.

Product Description

The OMSO2 product is written as HDF-EOS5 swath file. Data files are available from Goddard Earth sciences Data and Information Services Center (GES DISC) web site. For a list of tools that read HDF-EOS5 data files, please visit this link:
http://disc.gsfc.nasa.gov/Aura/tools.shtml

A file, also called a granule, contains SO_2 and associated information retrieved from each OMI scene from the sun-lit portion of an Aura orbit. The data are ordered in time sequence. The information provided on these files includes: latitude, longitude, solar zenith angle, OMTO3 reflectivity (LER) and independent estimates of the SO_2 vertical columns, as a well as a number of ancillary parameters that provide information to assess data quality. Four values of SO_2 column amounts are provided corresponding to four assumed vertical profiles. Independent information is needed to decide which value is most applicable. For a complete list of the parameters, please read the OMSO2 file Specification.

For general assistance with data archive, please contact GES DISC. For questions and comments related to the OMSO2 algorithm and data quality please contact Nickolay Krotkov (Nickolay.A.Krotkov@nasa.gov), who has the overall responsibility for this product, with copies to Can Li (Can.Li@nasa.gov).

The subsets of OMSO2 data over many ground stations and along Aura validation aircraft flights paths are also available through the Aura Validation Data Center (AVDC) web site.

References

Bhartia, P. K. and C. W. Wellemeyer (2002), OMI TOMS-V8 Total O3 Algorithm, *Algorithm Theoretical Baseline Document: OMI Ozone Products*, edited by P. K.

Bhartia, vol. II, ATBD-OMI-02, version 2.0. Available: http://eospso.gsfc.nasa.gov/eos_homepage/for_scientists/atbd/docs/OMI/ATBD-OMI-02.pdf

Carn, S. A., A. J. Krueger, N. A. Krotkov, K. Yang, and P. F. Levelt (2007a), Sulfur dioxide emissions from Peruvian copper smelters detected by the Ozone Monitoring Instrument, Geophys. Res. Lett., 34, L09801, doi:10.1029/2006GL029020.

Carn, S.A., N.A. Krotkov, K. Yang, R.M. Hoff, A.J. Prata, A.J. Krueger, S.C. Loughlin, and P.F. Levelt (2007b), Extended observations of volcanic SO_2 and sulfate aerosol in the stratosphere, *Atmos. Chem. Phys. Discuss.*, 7, 2857-2871. (http://www.atmos-chem-phys-discuss.net/7/2857/2007/acpd-7-2857-2007.html)

Carn, S.A., A.J. Krueger, N.A. Krotkov, S. Arellano, and K. Yang (2008a), Daily monitoring of Ecuadorian volcanic degassing from space, *J. Volcanol. Geotherm. Res.,* (in press).
Carn, S.A., A.J. Krueger, N.A. Krotkov, K. Yang, and K. Evans (2008b), Tracking volcanic sulfur dioxide clouds for aviation hazard mitigation. *Natural Hazards, Special Issue on Aviation Hazards from Volcanoes* (in press).

Joiner, J., L. Guanter, R. Lindstrot, M. Voigt, A. P. Vasilkov, E. M. Middleton, K. F. Huemmrich, Y. Yoshida, and C. Frankenberg (2013), Global monitoring of terrestrial chlorophyll fluorescence from moderate-spectral-resolution near-infrared satellite measurements: methodology, simulations, and application to GOME-2, *Atmos. Meas. Tech.*, 6, 2803-2823, doi:10.5194/amt-6-2803-2013.

Krotkov,N.A., B. McClure, R. Dickerson, S. Carn, Can Li, P.K. Bhartia, K. Yang, A. Krueger, Z. Li, P. Levelt, H. Chen, P.Wang, and D. Lu (2008), Ozone Monitoring Instrument (OMI) SO_2 validation over NE China, *J. Geophys. Res*., *Aura validation special issue*, (in press)

Krotkov, N.A., S.A. Carn, A.J. Krueger, P.K. Bhartia, and K. Yang (2006). Band residual difference algorithm for retrieval of SO_2 from the Aura Ozone Monitoring Instrument (OMI). *IEEE Trans.Geosci. Remote Sensing*, *AURA special issue*, 44(5), 1259-1266, doi:10.1109/TGRS.2005.861932, 2006

Krueger, A.J., L.S. Walter, P.K. Bhartia, C.C. Schnetzler, N.A. Krotkov, I. Sprod, and G.J.S. Bluth (1995) Volcanic sulfur dioxide measurements from the total ozone mapping spectrometer instruments. *J. Geophys. Res*., 100(D7), 14057-14076, 10.1029/95JD01222.

Li, C., J. Joiner, N. A. Krotkov, and P. K. Bhartia (2013), A fast and sensitive new satellite SO_2 retrieval algorithm based on principal component analysis: Application to the ozone monitoring instrument, *Geophys*. *Res*. *Lett*., 40, doi:10.1002/2013GL058134.

Bogumil, K., J. Orphal, T. Homann, S. Voigt, P. Spietz, O.C. Fleischmann, A. Vogel, M. Hartmann, H. Kromminga, H. Bovensmann, J. Frerick, J.P. Burrows (2003),

Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: instrument characterization and reference data for atmospheric remote-sensing in the 230-2380nm region, Journal of Photochemistry and Photobiology, A:Chemistry, 157 , 167-184.

Yang, K., N. Krotkov, A. Krueger, S. Carn, P. K. Bhartia, and P. Levelt (2007), Retrieval of Large Volcanic SO_2 columns from the Aura Ozone Monitoring Instrument (OMI): Comparisons and Limitations, *J. Geophys. Res*., 112, D24S43, doi:10.1029/2007JD008825